

KRISTIAN LEWIS i JOSIP MIHALJEVIĆ*



Odostražni rječnik – što je, kako ga izraditi i čemu služi²

Ovisno o razredbenome kriteriju rječnici se dijele na različite načine. Prema obuhvaćenome leksiku rječnici se dijele na opće i posebne. Opći rječnici obuhvaćaju riječi koje pripadaju općemu jeziku te najčešće nazive. Posebni rječnici obuhvaćaju riječi ograničenoga područja (npr. terminološki rječnici obuhvaćaju samo nazive određene struke), određenoga funkcionalnog stila (npr. rječnik žargona), riječi samo određenoga podrijetla (rječnik stranih riječi), riječi među kojima postoji određeni odnos (npr. rječnik sinonima i antonima, rječnik homonima).

S obzirom na broj jezika koje obuhvaćaju, rječnici se dijele na jednojezične, dvojezične i višejezične. Građa je u rječniku raspoređena u rječničke ili leksikografske članke. S obzirom na odnos prema korpusu rječnici se dijele na napametne i korpusne. Korpusni rječnici dijele se na korpusno utemeljene rječnike (*corpus based*, obrađivač se služi korpusom, ali može slobodno procijeniti što treba unijeti u rječnik te rječnik može po potrebi dopunjavati) i korpusno vođene rječnike (*corpus driven*, u rječniku mora biti sve što se nalazi u korpusu i ne smije biti ništa čega u korpusu nema). S obzirom na razdoblje koje obuhvaćaju, rječnici se dijele na povijesne i suvremene. S obzirom na raspored građe rječnici se dijele na abecedne, sustavne i čestotne.

U abecednim se rječnicima natuknice nižu abecednim redom. Postoje rječnici u kojima su riječi raspoređene tako da su abecedirane od kraja, a ne od početka riječi. Takvi se rječnici nazivaju odostražnim (ili inverznim) rječnicima. Oni su veoma korisni za znanstvena istraživanja, npr. za proučavanje tvorbenih odnosa među riječima, ali mogu se upotrijebiti i pri pisanju pjesama (za pronalaženje rime), pri ispunjavanju križaljke i slično. Mnogi europski jezici imaju odostražne rječnike, pokoji od njih i više od jednoga. Takve rječnike imaju, primjerice, engleski, francuski, njemački i talijanski jezik, a od slavenskih jezika ruski, slovenski, ukrajinski i srpski. Hrvatski je jezik dosad bio zastupljen u rječniku Josipa Matešića *Rückläufiges Wörterbuch des Serbokroatischen* (1965. – 1967.), koji je mrežno dostupan na stranicama Sveučilišta u Innsbrucku na adresi <https://www.uibk.ac.at/slawistik/institut/matesic.html>. S obzirom na riječi koje obuhvaćaju, odostražni rječnici mogu biti opći i posebni. Opći odostražni rječnici obuhvaćaju opći leksički fond pojedinoga jezika, poput općih jednojezičnih rječnika, ali građu donose, kako smo već napomenuli, abecedirano od kraja riječi. Posebni odostražni rječnici

* Josip Mihaljević nastavnik je informatike u Školi za primalje Zagreb.

² Rad je izrađen na istraživačkome projektu *Hrvatski mrežni rječnik – Mrežnik* (IP-2016-06-2141), koji u cijelosti financira Hrvatska zaklada za znanost i koji se provodi na ustanovi nositelju projekta Institutu za hrvatski jezik i jezikoslovje.

donose građu jednoga područja ili nekoliko njih, a od rječnika te vrste u hrvatskome jeziku izdvojit ćemo odostražni rječnik koji su izradili Božidar Finka i Antun Šojat u radu *Govor otoka Žirja* iz 1968. Taj je rječnik obuhvatio specifične mjesne riječi govora otoka Žirja (zemljoradnja, stočarstvo, ribarstvo, domaćinstvo, društveni i duhovni život itd.) i onomastički fond (toponomastiku i antroponimiju Žirja), a dostupan je na adresi <https://hrcak.srce.hr/68849>. Martina Grčević objavila je 2008. *Imena hrvatskih naselja: abecedni i odostražni rječnik*, koji je također mrežno dostupan na adresi https://bib.irb.hr/datoteka/388851.Imena_hrv_naselja_2008_crosbi.pdf. Taj odostražni rječnik donosi, pak, sva imena suvremenih naselja u Hrvatskoj iz popisa Državnoga zavoda za statistiku. Na stranici projekta *Muško i žensko u hrvatskome jeziku* (ihjj.hr/projekt/musko-i-zensko-u-hrvatskome-jeziku/72/) nalazi se demoinačica općega odostražnog rječnika i demoinačica posebnoga odostražnog rječnika mocijskih parnjaka. Obje su demoinačice rječnika jednojezične, suvremene i korpusno utemeljene.

U sklopu toga projekta proučavaju se različiti aspekti muško-ženskih odnosa u hrvatskome jeziku (tvorbeni, leksikografski, sociolingvistički, pragmatički itd.). Aktualnost je toga problema očita i iz velikoga broja savjeta koji se odnose na tvorbu ženskih parnjaka od muških, ali i muških od ženskih, na način oslovljavanja muških i ženskih osoba, pisanje zakonskih tekstova u kojima se spominju muške i ženske osobe i slično. Problem je u najvećem mjeri povezan s mocijskom tvorbom, ali je katkad i širi od nje jer se muško-ženski odnosi ne mogu uvijek izraziti tvorbenim odnosom (*medicinska sestra – medicinski tehničar*). Pri analizi tvorbenih mogućnosti iznimno je korisno imati odostražni rječnik kako bi se moglo lako pretražiti sve riječi koje imaju određeni završetak (npr. *-ica, -ač, -inja, -lac*). Stoga se na stranici projekta nalazi opći odostražni rječnik, koji se temelji na radnome

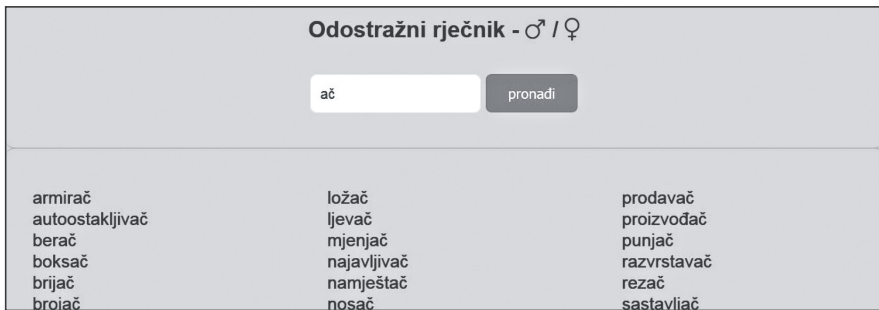
abecedariju *Hrvatskoga mrežnog rječnika – Mrežnika*, i specijalizirani rječnik mocijskih parnjaka, koji se temelji na radnome abecedariju muško-ženskih parova izrađenome u okviru projekta *Muško i žensko u hrvatskome jeziku*.

Rječnici su izrađeni s pomoću HTML5 i jQuery tehnologije. Početni kod za izradu rječnika preuzet je sa stranice Codepen (codepen.io/). Na toj su stranici dostupni različiti besplatni razvojni

Odostražni rječnik

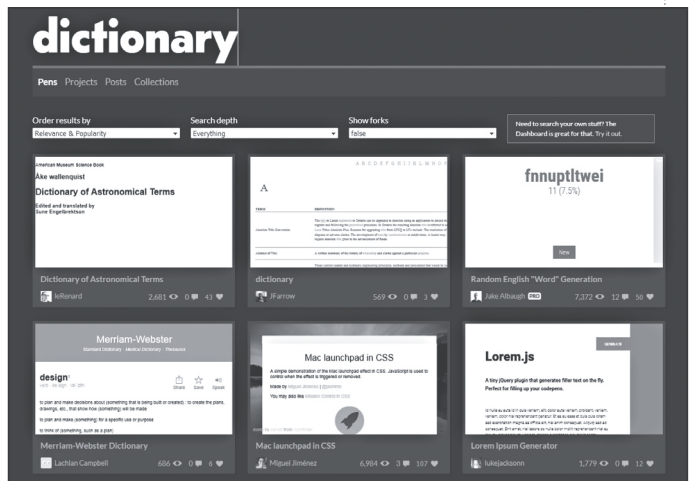
bacačičin beračičin biračičin boksačičin brijačičin crtičičin dizačičin djevojčičin dobavljačičin dostavljačičin	najavljičičin napadačičin navijačičin neplivačičin nepušačičin nečitačičin nosačičin objičičin ocjenjičičin opisivačičin
---------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------

1. slika: Opći odostražni rječnik



2. slika: Specijalni odostražni rječnik mocijskih parnjaka

kodovi za mrežne stranice, pa nije potrebno napredno znanje programiranja kako bi se iz bilo kojega popisa riječi izradili rječnici jer je logika pretraživanja i prikaza riječi ugrađena u razvojnim kodovima. U razvojnome kodu zamijenjen je popis riječi te dizajn i funkcionalnosti sučelja. Razvojni kod na toj stranici u početku nije bio razvijen za odostražno pretraživanje, ali je poslije to izmijenjeno. Pretraživanje odostražnoga rječnika provodi se s pomoću regularnih izraza ili skraćeno regexa³ koji učitavaju riječi iz JSON datoteke.⁴ Rječnici su responzivni, tj. mogu se pregledavati na svim računalnim uređajima te na različitim mrežnim preglednicima.



3. slika: Razvojni kodovi za rječnike na stranici Codepen

Rječnici imaju minimalistički dizajn u kojemu se jedino nalazi tražilica s pomoću koje se traže riječi. Kad se u tražilicu napiše nekoliko slova, dobiva se popis svih riječi iz

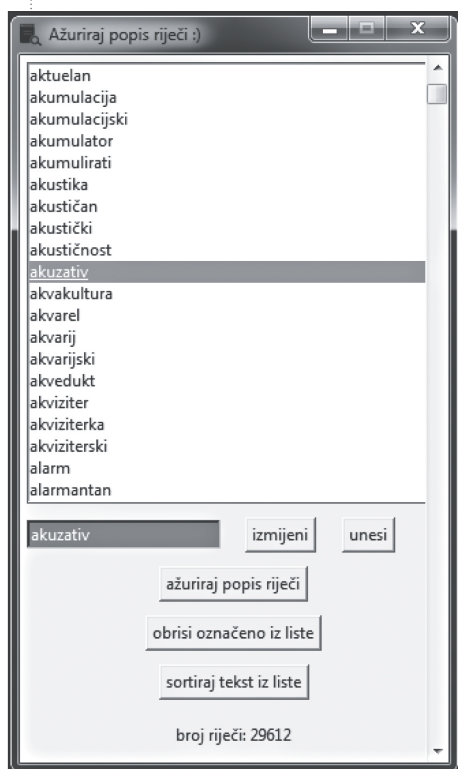
³ Regularni izrazi zadani su nizovi znakova čija se kombinacija može koristiti za pronalazak određenih dijelova teksta.

⁴ Datoteka JSON ili JavaScript Object Notation omogućuje organiziranu pohranu informacija u zagradama, a informacije se poslije mogu lako i brzo učitati na mrežnim stranicama. Datoteke JSON mogu se u nekim slučajevima upotrijebiti kao zamjena za XML datoteke.

početnoga popisa koje završavaju na ta slova. Kako je riječ o demoinačicama, zasad se poruke s pitanjima i prijedlozima mogu slati na adresu savjetnik@ihj.hr. Tako korisnici rječnika mogu sudjelovati u njihovoj izgradnji i širenju. Primjer je to masovne podrške (*crowdsourcing*), koja je sve popularnija u leksikografskome radu.

U okviru projekta *Hrvatski mrežni rječnik – Mrežnik* predviđena je izrada odostražnoga rječnika u četvrtoj (zadnjoj) godini projekta jer se završna inačica odostražnoga rječnika može izraditi tek nakon dobivanja završne inačice abecedarija. U programskome jeziku Python 3, koji je dostupan za besplatnu uporabu (python.org/), napravljeno je grafičko

sučelje koje omogućuje korisniku da pregleda popis riječi iz rječnika koje su zapisane u JSON datoteci. Korisnik može s pomoću grafičkoga sučelja, koje pokreće pritiskom na `.exe`⁵ datoteku, preuzeti sa stranice najnoviju inačicu popisa riječi, dodati nove riječi te izbrisati ili izmijeniti postojeće riječi. Izmjene se zapisuju u novu JSON datoteku, koja se može presnimiti preko stare JSON datoteke, koja se nalazi na mrežnoj stranici rječnika. Time autori rječnika mogu lako mijenjati i dopunjavati sadržaj rječnika te ne ovise o informatičaru. Stari popis riječi ili stara JSON datoteka pohranjuje se lokalno na računalu kao osiguranje dođe li do neželjenih izmjena. Na taj će se način nakon izrade konačne inačice abecedarija za *Mrežnik* s pomoću toga abecedarija lako izraditi odostražni rječnik *Mrežnika*.



4. slika: Grafičko sučelje izrađeno u Pythonu za izmjenu popisa riječi

⁵ Kratica za izvršnu datoteku koja pokreće program.